

# Using conversational chatbots to develop oral fluency in A1-level english learners.

## *Uso de chatbots conversacionales para desarrollar la fluidez oral en estudiantes de inglés de nivel A1.*

Vergara-Monrroy, Doris Ivonne<sup>1</sup>; Campoverde-López, Johnny Segundo<sup>2</sup>; Bonilla-Tenesaca, Josue Reinaldo<sup>3</sup>.

<sup>1</sup> Universidad Bolivariana del Ecuador; Ecuador, Duran; <https://orcid.org/0009-0007-7586-4291>; [divergaram@ube.edu.ec](mailto:divergaram@ube.edu.ec)

<sup>2</sup> Universidad Bolivariana del Ecuador; Ecuador, Duran; <https://orcid.org/0000-0003-0108-4755>; [jscampoverdel@ube.edu.ec](mailto:jscampoverdel@ube.edu.ec)

<sup>3</sup> Universidad Bolivariana del Ecuador; Ecuador, Duran; <https://orcid.org/0000-0002-6748-2345>; [jrbonillat@ube.edu.ec](mailto:jrbonillat@ube.edu.ec)

<sup>1</sup> Autor Correspondencia

 <https://doi.org/10.63618/omd/isj/v3/nE1/191>

**Cita:** Vergara-Monrroy, D. I., Campoverde-López, J. S., & Bonilla-Tenesaca, J. R. (2025). Using conversational chatbots to develop oral fluency in A1-level english learners. *Innova Science Journal*, 3(E1), 234-252. <https://doi.org/10.63618/omd/isj/v3/nE1/191>

**Recibido:** 27/08/2025

**Aceptado:** 05/12/2025

**Publicado:** 31/12/2025



**Copyright:** © 2025 por los autores. Este artículo es un artículo de acceso abierto distribuido bajo los términos y condiciones de la [Licencia Creative Commons, Atribución-NoComercial 4.0 Internacional. \(CC BY-NC\)](https://creativecommons.org/licenses/by-nc/4.0/).

(<https://creativecommons.org/licenses/by-nc/4.0/>)

**Resumen:** The teaching of oral expression in English at beginner levels is limited by a lack of authentic interaction, especially in Latin American contexts. This study evaluated the use of conversational chatbots to improve oral fluency in A1-level students in Ecuador, using a quasi-experimental design with pre- and post-tests in two groups (n = 70): an experimental group that worked with ABP and Microsoft Copilot, and a control group that received expository teaching. Fluency was measured using a CEFR-based rubric, and normality, t, Wilcoxon, Mann-Whitney, and Hake's gain tests were applied. The experimental group showed superior improvements and a moderate gain (g = 0.57), compared to the low progress of the control group (g = 0.23). It is concluded that integrating a conversational chatbot into PBL enhances oral fluency and communicative self-efficacy in A1 students.

**Palabras clave:** Artificial intelligence; project-based learning; English

**Abstract:** La enseñanza de la expresión oral en inglés en niveles iniciales se ve limitada por la escasa interacción auténtica, especialmente en contextos latinoamericanos. Este estudio evaluó el uso de chatbots conversacionales para mejorar la fluidez oral en estudiantes de nivel A1 en Ecuador, mediante un diseño cuasi-experimental con pretest y postest en dos grupos (n = 70): uno experimental que trabajó con ABP y Microsoft Copilot, y otro control con enseñanza expositiva. La fluidez se midió con una rúbrica basada en el MCER y se aplicaron pruebas de normalidad, t, Wilcoxon, Mann-Whitney y la ganancia de Hake. El grupo experimental mostró mejoras superiores y una ganancia moderada (g = 0,57), frente al progreso bajo del grupo control (g = 0,23). Se concluye que integrar un chatbot conversacional en el ABP potencia la fluidez oral y la autoeficacia comunicativa en estudiantes A1.

**Keywords:** Inteligencia artificial; aprendizaje basado en proyectos; inglés

## 1. Introducción

Learning English in basic education has become increasingly important in a world where intercultural communication and global skills are essential for student development. According to Tai & Chen (2024), developing oral expression at an early stage promotes not only communicative competence but also active student participation in broader academic and social environments. However, in Spanish-speaking contexts, particularly at initial levels such as A1, providing sufficient opportunities to practice the language remains a challenge due to the lack of real interaction and the limited time available in the classroom (Koç & Savaş, 2024).

Given these limitations, artificial intelligence-based technologies have emerged as a promising pedagogical alternative, especially for supporting oral practice and immediate feedback in foreign language learning. Recent evidence indicates that conversational chatbots can generate safe, personalized, and continuous practice environments, with positive effects on fluency, motivation, and reduced anxiety when speaking (Du, 2024; Lyu, 2025). However, studies agree that these resources reach their true potential when integrated into active methodologies that place the student in a leading role and encourage authentic language use within meaningful tasks (Tai & Chen, 2024).

The relevance of these tools is heightened in Latin America, where a lack of technological resources, large class sizes, and limited exposure to English outside the classroom hinder the sustained development of oral skills (UNESCO Office Santiago and Regional Bureau for Education in Latin America and the Caribbean et al., 2022). In Ecuador, especially in medium-sized cities such as Quevedo, these limitations are reflected in digital access gaps, unequal educational infrastructure, and limited opportunities for exposure to the language, which restricts the actual use of English in everyday contexts (Ministerio de Educación del Ecuador, 2022). In this regard, the literature emphasizes that integrating conversational technologies into active methodologies can help compensate for these barriers by providing a space for constant, accessible interaction geared toward meaningful communication (Koç & Savaş, 2024; Tai & Chen, 2024).

Within this framework, the present study aims to use conversational chatbots to develop oral fluency in seventh-grade A1-level English students at the Tungurahua Primary School in the parish of San Carlos, Quevedo canton.

## 2. Materiales y Métodos

### 2.1. Location

The study was conducted at Tungurahua Primary School, a public institution located in the parish of San Carlos, Quevedo canton, Los Ríos province, Ecuador (Figure 1). This educational center serves middle school students and has two classrooms for seventh grade, where the Project-Based Learning (PBL) teaching intervention was implemented, supported by a free conversational chatbot.

The classrooms have internet connectivity, institutional mobile devices, and adequate spaces for collaborative activities, which facilitated the integration of digital tools such as Microsoft Copilot, a resource selected as a free conversational chatbot for oral practice of English at level A1.

Figure 1.

*Location map of the study area*

**Note:** Geographic location of Tungurahua Primary School, San Carlos parish, Quevedo canton.  
**Source:** Own elaboration.

## 2.2. Type of research

The study was conducted using a quantitative approach, employing a quasi-experimental design with pre- and post-tests administered to two groups of students. This design allowed for a comparative analysis of the effect of using a conversational chatbot integrated into the Project-Based Learning (PBL) methodology on the development of oral fluency in students in the experimental group, in contrast to the performance achieved by the control group, which worked under a traditional expository methodology.

## 2.3. Research methods

- Descriptive method: this allowed the initial level of oral fluency (A1) to be characterized using a rubric based on the CEFR criteria (pronunciation, coherence, lexical range, and speech rate).
- Experimental method: this consisted of implementing PBL combined with a chatbot (Microsoft Copilot) in the experimental group (EG). In contrast, the control group (CG) worked with traditional reading and repetition activities. The

comparison between the two groups allowed us to determine the impact of the chatbot on the development of oral fluency.

**2.4. Population and sample**

The population consisted of 70 seventh-grade students from Tungurahua Elementary School. The sample corresponded to the two existing parallel classes, selected under a non-probabilistic criterion for convenience:

- Experimental Group (EG): 35 students, who worked with PBL + Microsoft Copilot chatbot.
- Control Group (CG): 35 students, who received traditional instruction.

Both groups had equivalent characteristics in terms of age, educational level, and prior proficiency in English.

**2.5. Research design**

The research was structured in four general phases: initial diagnosis, pedagogical intervention, subsequent evaluation, and statistical analysis of the results. Two naturally formed groups were used. In the experimental group, the pedagogical intervention was developed through a proposal that used Project-Based Learning integrated with the Microsoft Copilot chatbot as the main resource for oral practice. In the control group, the intervention followed a traditional teaching methodology focused on reading, repetition, and structured practice without technological mediation. This organization allowed for a comparison of the impact of both approaches on the development of students' oral fluency.

**2.5.1 Phase 1: Learning assessment**

In the initial phase of the study, an oral pretest was administered to all students in order to identify their level of communicative fluency in English prior to the intervention. The assessment consisted of a brief individual interaction designed for elementary school students, in which participants had to answer simple questions related to greetings, personal information, daily activities, and basic expressions relevant to level A1 (Table 1). The structure of the instrument was based on the descriptors of the Common European Framework of Reference for Languages (CEFR), which states that learners at level A1 should be able to produce simple sentences, recognize everyday expressions, and manage very brief communicative exchanges (Council of Europe, 2025).

**Table 1.**

**A1 A1 oral pre-test**

Section	Questions (level a1)
1. Greetings and introductions	1. Hello! What is your name? 2. How old are you? 3. Where are you from? 4. What grade are you in?
2. Personal information	5. Who do you live with? 6. What is your favorite color? 7. What is your favorite food? 8. Do you have brothers or sisters?

<p><b>3. Daily activities</b></p>	<p>9. What time do you wake up?                  10. What do you do after school?                  11. What time do you go to bed?                  12. What do you do on weekends?</p>
<p><b>4. Tastes and preferences</b></p>	<p>13. Do you like sports?                  14. What is your favorite animal?                  15. Do you like music?                  16. What is your favorite game?</p>
<p><b>5. Basic functional expressions</b></p>	<p>17. How are you today?                  18. Can you spell your name?                  19. What is the weather like today?                  20. What day is it today?</p>

**Note:** The questions were designed in accordance with the oral production descriptors for level A1 of the Common European Framework of Reference for Languages (Council of Europe, 2025) and adapted to the context of primary school students.

When assessing oral production in foreign languages, the specialized literature agrees that evaluating each response in isolation is methodologically insufficient, since communicative competence is not only manifested in the accuracy of a specific response, but also in the simultaneous integration of features such as pronunciation, fluency, intelligibility, lexical range, and interaction (Ismailia, 2021; Fulcher, 2015; Agasøster, 2015). Assessing oral skills using an individual scoring system for each question tends to fragment skills that in reality operate interdependently, which can lead to distortions in the interpretation of performance, especially in beginner learners. For this reason, international organizations such as the Council of Europe (2025) recommend the use of analytical rubrics, as they allow for the capture of performance gradients and a more accurate description of the CEFR levels of achievement.

In line with these recommendations, this research used an analytical rubric designed specifically for A1-level learners, which incorporated five fundamental dimensions for assessing oral production in primary school students (7th grade): pronunciation, temporal fluency, message intelligibility, use of basic vocabulary, and communicative interaction. The selection of these indicators was based on the principles of the CEFR, which emphasizes the need to assess communicative performance in terms of clarity, appropriateness, and the ability to sustain simple exchanges in basic contexts (Council of Europe, 2025). Table 2 presents the rubric used with the expected oral performance on scale 4.

**Table 2.**

**Analytical rubric with expected performance on scale 4**

Assessed Dimension	Scale 4 – Expected Performance at A1
Pronunciation	Articulates simple words and phrases with clear and understandable pronunciation, showing minimal interference from the mother tongue.
Temporal Fluency	Produces short utterances with a steady rhythm and natural pauses, maintaining continuity in the message without relevant interruptions.
Message Intelligibility	The message is understood without the need for frequent repetitions; ideas are organized simply and coherently within the A1 level.

<b>Basic Vocabulary</b>	Appropriately uses basic and frequent vocabulary to refer to everyday topics, selecting words that are relevant to the communicative context.
<b>Communicative Interaction</b>	Responds quickly to direct questions, sustains brief exchanges, and provides simple information spontaneously when the interlocutor offers minimal support.

**Note:** This table shows only the maximum performance expected of seventh-grade students in Basic Education.

Each dimension was rated using an ordinal scale from 0 to 4 points, where:

- 0 = no performance
- 1 = very limited performance
- 2 = partially achieved performance
- 3 = adequate performance at A1 level
- 4 = solid performance at A1 level

This structure allowed for a maximum total score of 20 points for each student. The use of ordinal scales is widely recommended in the assessment of oral expression, as it allows for the capture of gradients of progress that are not evident through dichotomous assessments (Fulcher, 2015).

Fulcher (2015) points out that converting qualitative performance scales into numerical values is a legitimate practice in oral assessment research, provided that the proportional relationship between the observed performance and the score awarded is preserved. In this way, the rubric allows for both qualitative and quantitative assessment, providing a measure that is sensitive to students' actual progress and suitable for comparative purposes between the experimental group and the control group.

### 2.5.2. Phase 2: Educational intervention

The pedagogical intervention was developed using two clearly differentiated approaches. In the experimental group, Project-Based Learning (PBL) was implemented, a methodology that promotes the active construction of knowledge based on authentic tasks and meaningful products (Tapia et al., 2025). Within this approach, the use of the Microsoft Copilot conversational chatbot was integrated as a central element, employed as a support tool for oral practice (Microsoft, 2025). The literature indicates that educational chatbots promote communicative interaction, increase oral practice, and reduce student anxiety at initial levels (Segura et al., 2025). Along these lines, students used Microsoft Copilot to practice basic vocabulary, model simple structures, and receive immediate feedback as they progressed through the different phases of the project.

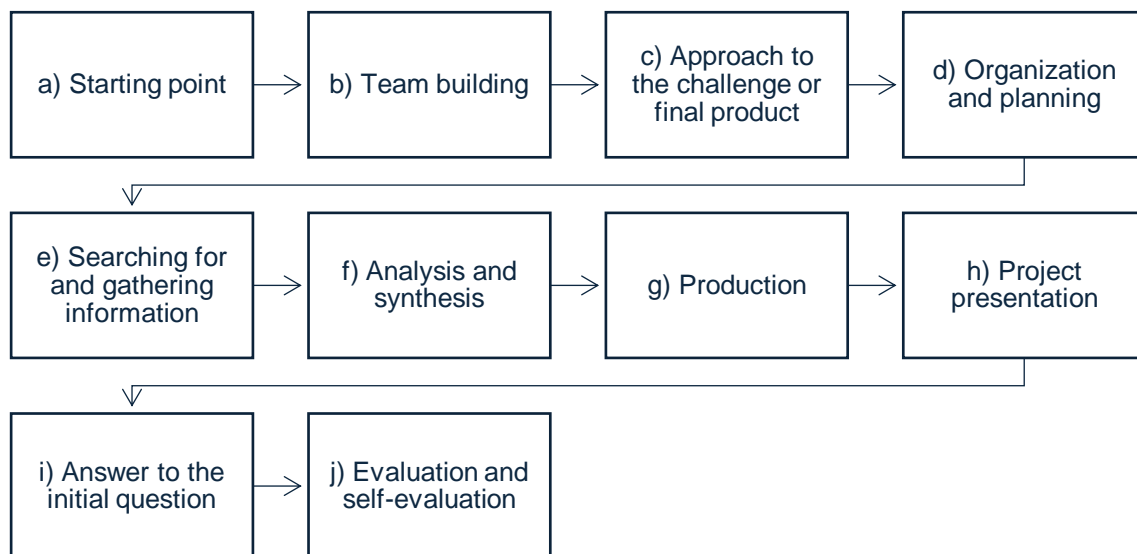
In contrast, the control group worked on the same A1-level content using a traditional methodology based on reading aloud, repetition, and structured exercises, an approach that continues to be common in language teaching in primary education (Richards & Rodgers, 2010). This modality did not include the use of technologies or PBL activities, which allowed for a comparison of the specific effect of using the chatbot on the development of oral fluency.

#### 2.5.2.1. Pedagogical intervention in the Experimental Group (EG): Use of Microsoft Copilot through Project-Based Learning

The pedagogical intervention in the experimental group was developed following the stages of Project-Based Learning (PBL) described by the Mexican Ministry of Public Education as a process organized into phases of exploration, planning, research, product construction, and socialization of results (Ministerio de Educación Pública, 2022). This framework was adapted to the context of teaching English at level A1 and complemented with the use of the Microsoft Copilot chatbot as a conversational resource to expand opportunities for oral practice. The PBL phases implemented in this group are described below (Figure 2):

**Figure 2.**

**Phases of Project-Based Learning**



**Nota:** Stages of project-based learning Source: Secretaría de Educación Pública de México

- a) At the outset, the teacher presented the following scenario: “Imagine you are participating in a school exchange with students from another country and you have to introduce yourself in English. What would you need to say so that they can get to know you better?” “Would you be able to introduce yourself properly in a real school exchange?” The students analyzed the exchange situation and reflected on what minimum information is necessary to introduce themselves to others in a clear and understandable way, considering only elements of level A1 of the CEFR. Simple examples of model dialogues were shown, and Microsoft Copilot was introduced as a conversational companion that would allow them to practice phrases, request rephrasing, rehearse pronunciation, and simulate mini-interactions.
- b) Teams of five members were then organized, seeking a balance in communication skills and confidence in speaking English. Each group reflected on the distribution of responsibilities, which allowed them to establish rotating roles that promoted collaboration and self-regulation, essential principles of PBL described by Krajcik & Shin (2014).

- c) The challenge was the central focus of the intervention. Each student had to produce an individual oral presentation lasting between 2 and 4 minutes that was clear, understandable, and consistent with the A1 level descriptors. However, the script or structure had to be worked on as a group. The presentation had to include basic greetings, name, age, country or city of origin, likes, and simple everyday activities. However, the methodological innovation lay in the fact that each part of the presentation had to be constructed and rehearsed through guided interactions with Microsoft Copilot. The chatbot was used to generate model sentences, practice simple structures, receive immediate feedback, compare alternative expressions, and simulate dialogues typical of an encounter between students from different countries. Thus, the development of the final product was not limited to producing a memorized text, but emerged from the process of linguistic negotiation between the student and the chatbot, which favored the development of fluency and exposure to comprehensible input according to the principles of communicative teaching (Nation & Newton, 2020).
- d) During the organization and planning phase, each group developed a detailed schedule that included conversation practice sessions with Copilot, vocabulary collection, progressive sentence construction, oral drafts, and guided essays. The teacher reviewed the schedules, offered feedback, and provided a guide with achievement criteria to guide the preparation of the final presentation.
- e) During the information search and collection phase, students consulted basic CEFR expressions, examples of short presentations, and phrases generated by Copilot. The chatbot was used to request simple explanations of vocabulary, obtain contextualized examples of usage, generate models of intercultural dialogues, and reformulate unclear statements. This stage allowed students to build a linguistic repertoire adapted to their real communication needs.
- f) Subsequently, in the analysis and synthesis stage, each team selected and organized the information gathered to turn it into meaningful content. Students compared versions of phrases suggested by Microsoft Copilot, identified the most appropriate expressions to introduce themselves, and adjusted the level of difficulty. The chatbot allowed them to verify the intelligibility of the message and adjust the presentation to ensure it was clear to a foreign listener.
- g) The production phase consisted of creating the video that included the oral presentation. Each group constructed their speech with the help of Copilot, requesting rephrasing, examples, corrections of common mistakes, and intonation models.
- h) In the project presentation phase, each student presented their video to the class. After sharing their final product (video), they participated in a brief spontaneous exchange with their classmates, where they answered simple questions similar to those they had practiced with Copilot. This moment allowed us to observe the transfer of the skills acquired to a real communicative situation in the classroom.
- i) The initial question response stage consisted of collectively reflecting on whether students would be able to introduce themselves appropriately in a real school exchange. Each student analyzed what elements they were able to incorporate, which phrases they handled with greater confidence, and how the Copilot exercises contributed to improving their clarity and fluency.

- j) Finally, the evaluation and self-evaluation phase was carried out, in which students completed an individual evaluation and also performed a self-evaluation that consisted of reflecting on the difficulties encountered during the learning process.

### **2.5.2.2. Educational intervention in the Control Group (CG): Implementation of a traditional lecture-based methodology**

The intervention applied to the control group was developed using a traditional expository methodology, characterized by the direct presentation of content and the structured practice of linguistic patterns. In this approach, the teacher leads the activities and the students perform predominantly receptive and reproductive tasks, in line with current descriptions of teacher-centered instruction (McLeod, 2024). The sessions followed a fixed sequence that included reading aloud, choral repetition, substitution exercises, and closed-ended activities typical of the A1 level, while the teacher recorded student participation in a field note.

The content covered greetings, personal introductions, everyday vocabulary, and basic grammatical structures. The activities were based exclusively on the textbook and printed guides, without the use of digital resources or technology-mediated interaction. This type of instruction is considered representative of direct teaching practices that prioritize the initial acquisition of the linguistic system through controlled exposure and guided practice (Nation & Newton, 2020).

### **2.5.3. Phase 3: Learning assessment**

The learning achieved by students in both groups (experimental group and control group) was evaluated by means of an oral post-test. In order to ensure equivalent conditions and allow for valid comparisons, the post-test was designed as an exact methodological replica of the pre-test. To this end, the same structured and categorized questions presented in Table 1 were used, corresponding to four communicative areas of level A1: greetings, personal information, daily activities, and tastes. The reuse of these same questions ensured the stability of the construct measured and allowed for the identification of changes attributable exclusively to the teaching intervention.

Similarly, the assessment of oral performance in the post-test was carried out using the same analytical rubric summarized in Table 2, consisting of five fundamental dimensions for level A1: pronunciation, temporal fluency, intelligibility of the message, use of elementary vocabulary, and communicative interaction. The use of a single instrument for both measurements ensured consistency in the evaluation criteria and intra-instrument reliability throughout the process.

#### **2.5.3.1. Phase 4: Statistical analysis of data**

The statistical treatment of the information was organized in several stages with the purpose of determining whether the intervention produced significant changes in the students' oral fluency. First, the pretest and posttest scores of both groups were subjected to a distribution check using the Shapiro–Wilk test, which allowed us to identify whether or not the data met the assumptions of normality (Luzuriaga et al., 2023). This review was essential to decide, in each case, whether to use parametric or nonparametric techniques in subsequent analyses.

When the results showed behavior close to normality, Student's t-tests were applied (Sánchez Turcios, 2015), differentiating between internal comparisons (before and after within the same group) and external comparisons (posttest between groups). On the other hand, if the data did not conform to normality, comparisons were made using alternative procedures without strict assumptions, specifically the Wilcoxon test for pre-post contrasts within the same group (Centro Estadístico de la Universidad del Azuay, 2024) and the Mann–Whitney U test to examine differences between the experimental group and the control group at the end of the intervention (Romero Saldaña, 2013).

In addition to statistical contrasts, the level of progress achieved was estimated using the gain factor, an indicator that allows the observed progress to be related to the learning margin available to each student. This index was calculated using the formula established to measure relative gain (Hake, 1998) and was then interpreted according to the ranges commonly used in educational research, distinguishing between low, moderate, and high progress.

This gain ( $g$ ) was obtained using the formula proposed by Hake (1998):

$$g = \frac{\text{Posttest (\%)} - \text{Pretest (\%)}}{100 - \text{Pretest (\%)}}$$

Where:

- Low ( $g \leq 0.3$ )
- Medium ( $0.3 < g \leq 0.7$ )
- High ( $g > 0.7$ ).

A gain greater than 0.3 represents a favorable indicator of learning achieved and the effectiveness of the methodology implemented.

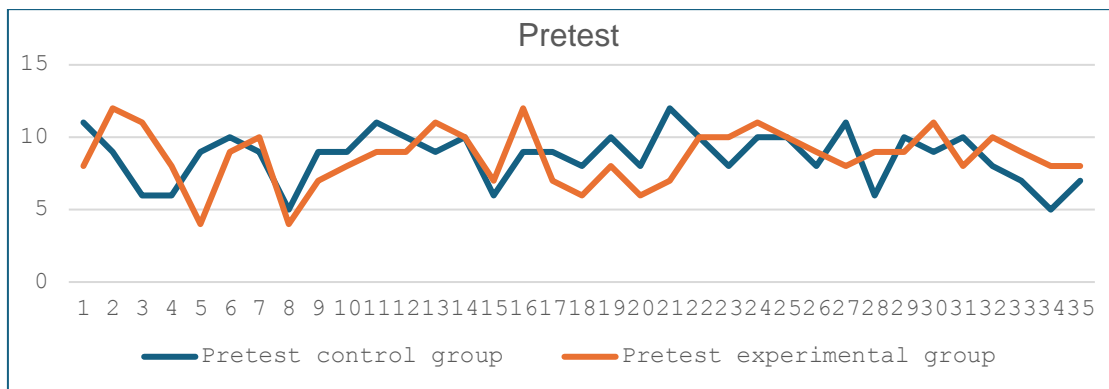
### 3. Resultados

#### 3.1. Phase 1: Learning assessment

The results showed that both groups started with equivalent levels of oral fluency, consistent with the A1 level expected for seventh-grade students. In quantitative terms, the Control Group obtained an average score of 8.69/20, while the Experimental Group achieved the same average (8.66/20), confirming that the initial conditions were comparable before the teaching intervention (Figure 3). This similarity ensures that any subsequent differences can reasonably be attributed to the instructional approaches applied during the study.

Figure 3.

**Comparison of pretest scores using multiple lines**



**Note:** The scores correspond to the oral pretest assessment administered to the 70 students using the CEFR A1 level analytical rubric.

**3.2. Phase 2: Educational intervention**

**3.2.1. Educational intervention in the experimental group**

During the intervention, students in the Experimental Group prepared their final oral presentations on video, which constitute the main product of the project. Analysis of these materials showed that most students were able to produce simple, coherent, and understandable speech in English, using complete sentences appropriate for level A1 and organizing information in a linear fashion (personal introduction, family, routines, and likes). The videos were within the expected range of 2 to 4 minutes in length, and students demonstrated appropriate use of the basic vocabulary learned during the sessions. The products reflected early qualitative progress in sentence construction, basic vocabulary selection, and control of speaking pace. Likewise, greater confidence in oral expression was observed, especially among students who initially expressed fear or difficulty in speaking English. Table 3 below presents representative examples of the discourse produced by the seven groups in the Experimental Group.

Table 3.

**Representative excerpts from the final speech prepared by the seven groups of the Experimental Group**

Group	Representative speech (extended A1)
Group 1 student	“Hello, my name is Sofia. I am twelve years old. I am from Ecuador and I live in Quevedo with my parents. I wake up every day at six o'clock because my school starts early. After school I usually read, draw or help at home. I like listening to music and sometimes I play games with my friends. My favorite food is pasta and my favorite color is pink. On weekends I visit my cousins or watch movies. I like dogs because they are friendly. This is my presentation. Thank you.”
Group 2 student	“Good morning, my name is Mateo. I am twelve years old. I am from Ecuador and I live in the city of Quevedo with my family. I wake up early to get ready for school, and after school I sometimes play soccer or help my mother. My favorite food is rice with chicken and my favorite color is blue. On weekends I play with my brother or ride my bike. I also like watching cartoons. I have one brother and we study together. This is a little about me.”

**Group 3 student**

"Hello, my name is Daniela. I am twelve years old and I am from Ecuador. I live in Quevedo with my grandmother. Every day I wake up at six thirty. After school I like playing volleyball or talking with my friends. My favorite food is soup and my favorite color is purple. On weekends I visit my family or help at home. I also like reading short stories. My favorite animal is the cat because it is quiet and cute. Thank you for listening."

**Group 4 student**

"Hi, my name is Luis. I am twelve years old. I am from Ecuador and I live in Quevedo. I wake up at six in the morning. After school I do my homework and then I play soccer with my friends. My favorite food is chicken and my favorite color is green. On weekends I visit my grandparents or watch TV. I have one sister and she is younger than me. I like dogs and birds. This is my personal presentation."

**Group 5 student**

"Hello, my name is Karla. I am twelve years old and I am from Ecuador. I live in Quevedo with my parents and my sister. Every day I wake up at six o'clock, go to school, and after school I like drawing or listening to music. My favorite food is spaghetti and my favorite color is yellow. On weekends I play with my sister or visit my cousins. My favorite animal is the rabbit because it is soft and cute. This is my presentation about myself."

**Group 6 student**

"Hi everyone, my name is Andrés. I am twelve years old. I am from Ecuador and I live in Quevedo with my family. I wake up early to go to school. After school I help my mother, read a little or play soccer. My favorite food is fish and my favorite color is red. On weekends I play with my friends or visit my grandparents. My favorite animal is the dog. I also like drawing in my notebook. Thank you."

**Group 7 student**

"Good afternoon, my name is María. I am twelve years old. I am from Ecuador and I live in Quevedo with my parents and my two brothers. Every day I wake up at six o'clock, go to school, and after school I study or help my mother. My favorite food is chicken soup and my favorite color is blue. On weekends I play with my brothers or watch movies. My favorite animal is the bird. This is my short presentation."

**Note:** The speeches were selected as representative samples of the performance achieved by students during the oral post-test at level A1.

At the end of the project, the students reflected on the question posed at the beginning: "Would you be able to introduce yourself properly in a real school exchange?" The general response was affirmative. The students said that they did feel capable of introducing themselves in English, indicating that they can now:

- Give basic information about themselves
- Follow a logical order when speaking
- Answer simple questions
- Use common phrases without relying entirely on memory
- Speak with greater confidence after practicing with the chatbot

Several students said that before the project, they "didn't know what to say" or "felt embarrassed," but thanks to repeated practice with Microsoft Copilot, they were able to better structure their ideas and feel more confident when speaking English.

This reflection shows that the project not only allowed for the creation of a final product, but also increased the students' perception of communicative self-efficacy, which is a

relevant qualitative result prior to the measurement of formal learning reported in the next phase.

**3.2.2. Educational intervention in the control group**

The data collected shows that, as the weeks progressed, most students improved their accuracy in repeating short phrases and identifying basic vocabulary. This progress was reflected in a gradual reduction in errors when pronouncing simple greetings, responding with personal information (name, age, origin), and using very frequent expressions at the A1 level. Likewise, it was observed that a growing number of students were able to answer closed questions without long pauses, especially those practiced repeatedly during classroom activities.

However, the records also show that the group had limitations in tasks that required more spontaneous production. The teacher’s notes describe how students tended to memorize short structures without extending their responses, and that prolonged pauses and word searching persisted in questions related to everyday activities or personal tastes.

**3.4. Phase 3 (Learning assessment) and phase 4 (Statistical analysis of data)**

The Shapiro–Wilk test was used to verify whether the pretest and posttest scores followed a normal distribution and, therefore, whether it was appropriate to use parametric tests. The results showed that three of the four measurements (pretest and posttest of the experimental group and posttest of the control group) had p-values greater than 0.05, indicating behavior consistent with normality. The only exception was the control group's pretest, whose p-value = 0.0306 showed a non-normal distribution (Table 4). This difference made it necessary to select different tests for each comparison: when both measurements were normal, parametric tests (Student's t-test) were used, while when either of them was not normal, nonparametric tests (Wilcoxon or Mann–Whitney) were used.

**Table 4.**

***Representative excerpts from the final speech prepared by the seven groups of the Experimental Group***

Group	Time point	p-value	Interpretation
Control	Pretest	0.0306	Not normal ( $p < 0.05$ )
Control	Posttest	0.2417	Normal ( $p > 0.05$ )
Experimental	Pretest	0.1291	Normal ( $p > 0.05$ )
Experimental	Posttest	0.2118	Normal ( $p > 0.05$ )

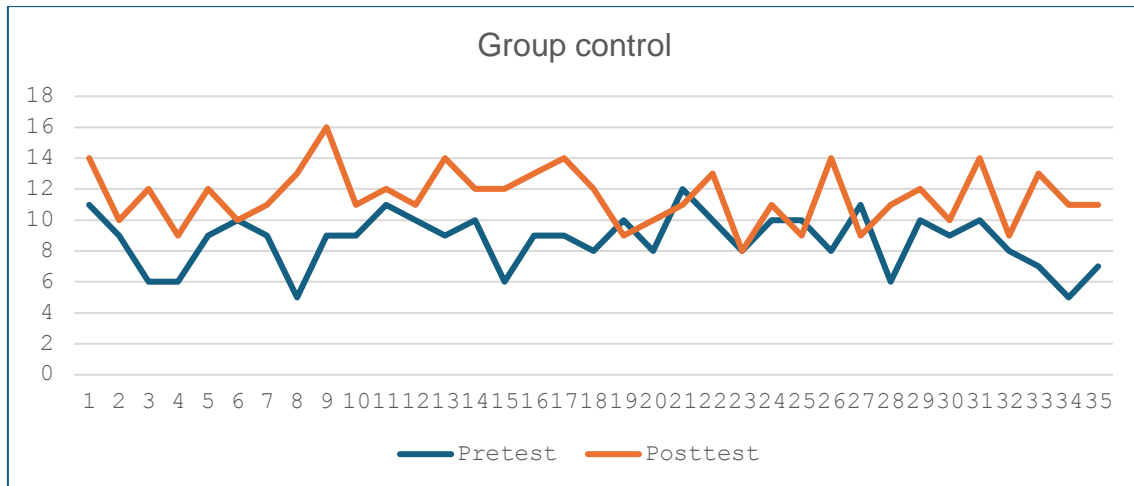
**Note:**  $p < 0.05$  indicates non-normal distribution;  $p > 0.05$  indicates distribution compatible with normality according to Shapiro–Wilk.

Based on the above, the internal comparison of the Control Group between the pretest and posttest was performed using the Wilcoxon test, which showed a statistically significant change ( $W = 27.5$ ;  $p < 0.001$ ). This indicates that, although the students

worked under the traditional expository methodology, they did achieve a significant, albeit moderate, improvement in their oral production (Figure 4).

**Figure 4.**

**Performance in pretest and posttest of the control group**

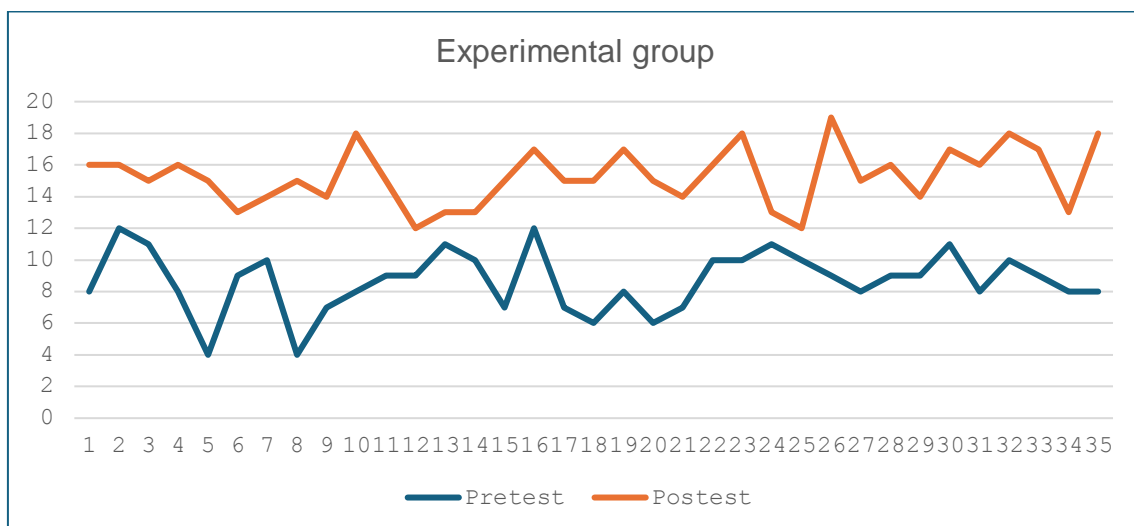


**Note:** The graphs show the individual comparison between pretest and posttest scores of the control group obtained using the analytical rubric for level A1.

In the Experimental Group, both measurements were normal, so a paired Student's t-test was applied. The results were highly significant ( $t(34) = -15.00$ ;  $p < 0.001$ ), showing that the intervention using chatbots through Project-Based Learning with intensive oral practice produced a much greater increase in students' oral performance (Figure 5).

**Figure 5.**

**Performance in pretest and posttest of the experimental group**



**Note:** The graphs show the individual comparison between pretest and posttest scores of the experimental group obtained using the analytical rubric for level A1.

To compare performance between groups, the pretest was evaluated first. Because one of the measurements was not normal, the Mann–Whitney U test was used, which

showed no significant differences between the groups at baseline ( $p > 0.05$ ). This confirms that both groups started under equivalent conditions, which is essential for attributing subsequent changes to the intervention.

Finally, the posttest comparison between groups, where both distributions were normal, was performed using Student's t-test for independent samples. The results revealed highly significant differences in favor of the Experimental Group ( $t(68) = -8.58$ ;  $p < 0.001$ ). In practical terms, this demonstrates that the pedagogical intervention implemented in that group produced clearly superior improvements to those obtained under the traditional methodology.

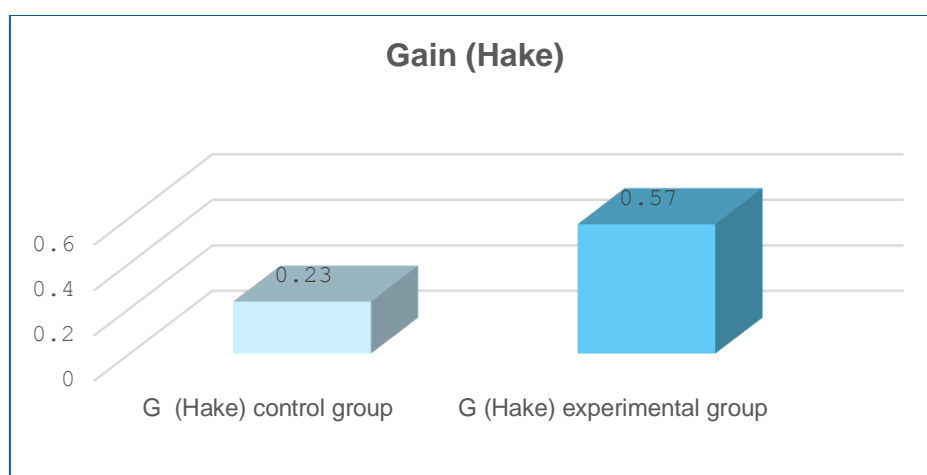
#### Hake's gain analysis showed clear differences between the groups:

- Control Group: obtained an average gain of  $g = 0.23$ , a value that falls within the low gain range. This indicates that, although there was a slight improvement between the pretest and posttest, students only achieved approximately 23% of the potential learning available. Progress was limited and consistent with the expected impact of a traditional methodology.
- Experimental Group: achieved an average gain of  $g = 0.57$ , corresponding to a moderate level of gain. This value implies that students recovered about 57% of the possible learning, reflecting considerably greater progress than that observed in the control group. The level of progress shows a favorable response to the intervention applied.

Taken together, these gains reinforce the previous statistical results and confirm that the strategy implemented in the experimental group produced broader, deeper, and more sustained improvements than those generated by the traditional expository methodology (Figure 6).

**Figure 6.**

***Gain calculated according to Hake 1998: In the control group and experimental group***



**Note:** The graphs show the comparison of the gain according to Hake 1998 in both groups

#### 4. Discusión

After implementing the PBL methodology with the use of chatbots, students were able to produce simple but structured speeches, meeting the requirements of level A1. This progress coincides with the findings of Gonzales-Ticona & Jacobs-Estrada (2025), who argue that AI increases efficiency in processes that require repetition, immediate feedback, and automation of simple tasks. In this sense, the chatbot functioned as a “conversational assistant” that allowed students to practice without social pressure, with constant feedback and at a personalized pace, which is usually difficult to achieve in classes with large groups.

The study conducted by Moreira-Vera & Lozano Alvarado (2025) shows that teachers agree that English acquisition is enhanced when students participate in communicative activities that encourage them to use the language naturally, without focusing excessively on formal correctness. Strategies such as storytelling, games, pair work, and projects promote this natural use of language. This view coincides with what the chatbot enables within PBL: by interacting with a partner who responds immediately, flexibly, and without generating evaluative pressure, students find a space that resembles the authentic communicative situations described by teachers. In this way, the tool not only reinforces content, but also enables conditions close to acquisition, as it promotes a freer and more meaningful use of English that transcends exclusively structured practice.

Internationally, recent research supports this interpretation and shows that the use of chatbots in language learning contexts brings direct benefits in terms of students' oral production and communicative confidence. For example, Tai and Chen (2024) show that, in primary education, interaction with generative chatbots allows for greater length and fluency in oral interventions, while reducing anxiety when speaking. Similarly, Tai & Chen (2024) reports significant improvements in willingness to communicate in English when children practice with chatbots that simulate real conversational situations. These findings are consistent with those observed in this study, as the chatbot acted as a constant interlocutor that sustained dialogue and offered opportunities for genuine practice, qualities that are difficult to replicate with traditional oral activities alone.

Likewise, recent reviews such as those by Guartán Guamán and Valdiviezo Ramírez (2025), Zhai and Wibowo (2023), and Li et al. (2025) highlight that chatbots are especially useful when incorporated into authentic tasks, simulations, and projects, as this promotes meaningful interaction rather than just mechanical repetition. These findings explain why combining PBL with the use of chatbots produced more solid progress: the digital tool was integrated into a communicative task with a real purpose, enhancing both practice and acquisition and placing the student in an active and dialogical role.

## 5. Conclusiones

The use of a conversational chatbot integrated into ABP produced clearly superior learning: the Experimental Group achieved a moderate gain ( $g = 0.57$ ), while the Control Group obtained a low gain ( $g = 0.23$ ). Both groups started from equivalent levels of oral fluency (pretest = 8/20 in both cases), so the final differences are attributed to the intervention with PBL + chatbot. At the intra-group level, the improvement was significant in both cases, but more intense in the Experimental Group ( $t(34) = -15.00$ ;  $p < 0.001$ ) than in the Control Group ( $W = 27.5$ ;  $p < 0.001$ ), showing a greater impact of the chatbot on pronunciation, fluency, intelligibility, vocabulary, and interaction. The post-test comparison between groups confirmed the superiority of the chatbot approach, with highly significant differences in favor of the Experimental Group ( $t(68) = -8.58$ ;  $p < 0.001$ ), supporting its use as an effective resource for developing oral fluency in A1 students. Thus, the incorporation of a conversational chatbot into Project-Based Learning (PBL) is consolidated as an effective pedagogical alternative for strengthening oral fluency in A1-level students.

## Referencias Bibliográficas

- Agasøster, S. (2015). Assessment of oral English: A study of assessment practice of oral English at lower secondary schools in Norway [Tesis de maestría, University of Bergen]. University of Bergen Repository. <https://files.core.ac.uk/download/pdf/30862951.pdf>
- Centro de Estadística, Universidad del Azuay. (2024, julio). Boletín investigativo No. 53: Herramientas estadísticas: prueba de Wilcoxon [Boletín]. Cuenca, Ecuador. <https://centro-estadistica.uazuay.edu.ec/sites/centro-estadistica.uazuay.edu.ec/files/2024-07/Bolet%C3%ADn%20Investigativo%20No.%2053.pdf>
- Council of Europe. (2025). Common European Framework of Reference for Languages (CEFR): Learning, teaching, assessment. <https://www.coe.int/en/web/common-european-framework-reference-languages>
- Du, J., & Daniel, B. K. (2024). Transforming language education: A systematic review of AI-powered chatbots for English as a foreign language speaking practice. *Computers and Education: Artificial Intelligence*, 6, 100230. <https://doi.org/10.1016/j.caeai.2024.100230>
- Fulcher, G. (2015). Re-examining language testing: A philosophical and social inquiry. Routledge. <https://doi.org/10.4324/9781315695518>
- Gonzales-Ticona, Y. B., & Jacobs-Estrada, M. C. (2025). La Inteligencia Artificial y su Impacto en la Brecha entre Países Desarrollados y en Vías de Desarrollo. *Innova Science Journal*, 3(4), 626-640. <https://doi.org/10.63618/omd/isj/v3/n4/112>
- Guartán Guamán, S. P., & Valdiviezo Ramírez, E. A. (2025). Using chatbots as a tool to personalize english as a foreign language learning: A literature review. *Runas. Journal of Education and Culture*, 6(12), e250283. <https://doi.org/10.46652/runas.v6i12.283>

- Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, 66(1), 64–74. <https://doi.org/10.1119/1.18809>
- Ismailia, T. (2021). Performance assessment using rubric to improve students' speaking skill. *JALL: Journal of Applied Linguistics and Literacy*, 5(2), 178–188. <https://doi.org/10.33974/jall.v5i2.335>
- Krajcik, J. S., & Shin, N. (2014). Project-based learning. En R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 275–297). Cambridge University Press. <https://doi.org/10.1017/CBO9781139519526.018>
- Koç, F. Ş., & Savaş, P. (2024). The use of artificially intelligent chatbots in English language learning: A systematic meta-synthesis study. *ReCALL*, 37(1), 4–21. <https://doi.org/10.1017/S0958344024000168>
- Li, Y., Zhou, X., Yin, H., et al. Design language learning with artificial intelligence (AI) chatbots based on activity theory from a systematic review. *Smart Learn. Environ.* 12, 24 (2025). <https://doi.org/10.1186/s40561-025-00379-0>
- Lyu, B., Lai, C., & Guo, J. (2025). Effectiveness of chatbots in improving language learning: A meta-analysis of comparative studies. *International Journal of Applied Linguistics*, 35(2), 527–953. <https://doi.org/10.1111/ijal.12668>
- Luzuriaga Jaramillo, H. A., Espinosa Pinos, C. A., Haro Sarango, A. F., & Ortiz Román, H. D. (2023). Histograma y distribución normal: Shapiro-Wilk y Kolmogorov-Smirnov aplicado en SPSS. *LATAM Revista Latinoamericana de Ciencias Sociales y Humanidades*, 4(4), 596. <https://doi.org/10.56712/latam.v4i4.1242>
- McLeod, S. (2024, February 1). Enseñanza expositiva: Teoría del aprendizaje de Ausubel. *Simply Psychology*. <https://www.simplypsychology.org/teoria-del-aprendizaje-ausubel.html>
- Microsoft. (2025). What is Microsoft 365 Copilot? Microsoft Learn. <https://learn.microsoft.com/en-us/copilot/microsoft-365/microsoft-365-copilot-overview>
- Ministerio de Educación del Ecuador. (2022). Informe preliminar del 2021. Ministerio de Educación. [https://educacion.gob.ec/wp-content/uploads/downloads/2022/03/informe\\_narrativo\\_rendicion\\_cuentas\\_2021.pdf](https://educacion.gob.ec/wp-content/uploads/downloads/2022/03/informe_narrativo_rendicion_cuentas_2021.pdf)
- Moreira-Vera, C. A., & Lozano Alvarado, C. I. (2025). ¿Aprender o adquirir el inglés? Una reflexión crítica en una escuela pública ecuatoriana. *Innova Science Journal*, 3(4), 290303. <https://doi.org/10.63618/omd/isj/v3/n4/137>
- Nation, I. S. P., & Newton, J. (2020). *Teaching ESL/EFL listening and speaking* (2nd ed.). Routledge. <https://doi.org/10.4324/9780429203114>
- Richards, J. C., & Rodgers, T. S. (2010). *Approaches and methods in language teaching* (2nd ed.). Cambridge University Press. <https://doi.org/10.1017/CBO9780511667305>

- Romero Saldaña, M. (2013). Contraste de hipótesis: Comparación de dos medias independientes mediante pruebas no paramétricas: Prueba U de Mann-Whitney. *Enfermería del Trabajo*, 3, 77–84.
- Sánchez Turcios, R. A. (2015). t-Student. Usos y abusos. *Revista Médica del IMSS*, 26(1), 59–61. <https://www.scielo.org.mx/pdf/rmc/v26n1/v26n1a9.pdf>
- Secretaría de Educación Pública (SEP). (2022). Guía metodológica para la implementación del Aprendizaje Basado en Proyectos en la Educación Superior. Subsecretaría de Educación Superior. <https://educacionbasica.sep.gob.mx/wp-content/uploads/2022/06/Metodologia-ABP-Final.pdf>
- Segura Torres, M. J., Navas Espin, W. R., Zurita Hurtado, H. A., Parra López, R. A., & Zambrano Silva, D. H. (2025). Uso de chatbots inteligentes como herramienta estratégica para la mejora de la atención al cliente en empresas de telecomunicaciones en Ecuador. *Arandu*, 12(2), 3910. <https://doi.org/10.69639/arandu.v12i2.1198>
- Tai, T.-Y., & Chen, H. H.-J. (2024). Improving elementary EFL speaking skills with generative AI chatbots: Exploring individual and paired interactions. *Computers and Education*, 220. <https://doi.org/10.1016/j.compedu.2024.105112>
- Tapia, D., Freire, L., y Hallo, E., (2025). Aprendizaje Basado en Proyectos: Un enfoque educativo innovador para una enseñanza activa. *Reincisol*, 4(7), pp. 320-341. [https://doi.org/10.59282/reincisol.V4\(7\)320-](https://doi.org/10.59282/reincisol.V4(7)320-)
- UNESCO Office Santiago and Regional Bureau for Education in Latin America and the Caribbean, UNICEF Regional Office for Latin America and the Caribbean, & Comisión Económica para América Latina y el Caribe. (2022). Education in Latin America and the Caribbean at a crossroads: Regional monitoring report SDG 4 – Education 2030. UNESCO. <https://doi.org/10.54675/HOTK9971>
- Zhai, C., & Wibowo, S. (2023). A systematic review on artificial intelligence dialogue systems for enhancing English as foreign language students' interactional competence in the university. *Computers and Education: Artificial Intelligence*, 4, 100134. <https://doi.org/10.1016/j.caeai.2023.100134>

## CONFLICTO DE INTERESES

“Los autores declaran no tener ningún conflicto de intereses”.